UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

AUTHORS GUILD, DAVID BALDACCI, MARY BLY, MICHAEL CONNELLY, SYLVIA DAY, JONATHAN FRANZEN, JOHN GRISHAM, ELIN HILDERBRAND, CHRISTINA BAKER KLINE, MAYA SHANBHAG LANG, VICTOR LAVALLE, GEORGE R.R. MARTIN, JODI PICOULT, DOUGLAS PRESTON, ROXANA ROBINSON, GEORGE SAUNDERS, SCOTT TUROW, and RACHEL VAIL, individually and on behalf of others similarly situated,	Case No. 1:23-cv-08292-SHS
Plaintiffs,	
v.	
OPENAI INC., OPENAI LP, OPENAI LLC, OPENAI GP LLC, OPENAI OPCO LLC, OPENAI GLOBAL LLC, OAI CORPORATION LLC, OPENAI HOLDINGS LLC, OPENAI STARTUP FUND I LP, OPENAI STARTUP FUND GP I LLC, and OPENAI STARTUP FUND MANAGEMENT LLC,	
Defendants.	
JONATHAN ALTER, KAI BIRD, TAYLOR BRANCH, RICH COHEN, EUGENE LINDEN, DANIEL OKRENT, JULIAN SANCTON, HAMPTON SIDES, STACY SCHIFF, JAMES SHAPIRO, JIA TOLENTINO, and SIMON WINCHESTER, on behalf of themselves and all others similarly situated,	CASE NO. 1:23-CV-10211-SHS
Plaintiffs,	
v.	
OPENAI, INC., OPENAI GP, LLC, OPENAI, LLC, OPENAI OPCO LLC, OPENAI GLOBAL LCC, OAI CORPORATION, LLC, OPENAI HOLDINGS, LLC, and MICROSOFT CORPORATION,	
Defendants.	

JOINT STIPULATION AND [PROPOSED] ORDER RE
TRAINING DATA INSPECTION PROTOCOL

Upon the stipulation of the parties, the following protocol will apply to the inspection, review, and/or disclosure of Training Data produced by Defendants OpenAI, Inc., OpenAI, L.P., OpenAI L.L.C., OpenAI OpCo, L.L.C., OpenAI Global L.L.C., OAI Corporation L.L.C., OpenAI Holdings L.L.C., OpenAI GP, L.L.C., OpenAI Startup Fund GP I, L.L.C., OpenAI Startup Fund I, L.P., and OpenAI Startup Fund Management, LLC (collectively, "OpenAI"):

1.   For the purposes of this protocol, "Training Data" shall be defined as data used to train relevant OpenAI LLMs. OpenAI reserves the right to update the Training Data made available for inspection should it find additional responsive data.

2.   The "Inspecting Party" shall be defined as all plaintiffs collectively in the above captioned consolidated action, including their attorneys of record, agents, retained consultants, experts, and any other persons or organization over which they have direct control.

3.   Training Data shall be made available for inspection in electronic format at OpenAI's offices in San Francisco CA, or at a secure location determined by OpenAI. Training Data will be made available for inspection between the hours of 8:30 a.m. and 5:00 p.m. on business days, although the parties will be reasonable in accommodating reasonable requests to conduct inspections at other times.

4.   The Inspecting Party shall provide five days' notice prior to any inspection.

5.   Training Data shall be designated "HIGHLY CONFIDENTIAL – ATTORNEYS' EYES ONLY" pursuant to the Stipulated Protective Order entered in this case on May 16, 2024, and the Inspecting Party may disclose Training Data only to those authorized to view "HIGHLY CONFIDENTIAL – ATTORNEYS" EYES ONLY" information under paragraph 16 of the Stipulated Protective Order, without prejudice to any party's right to challenge this confidentiality

designation (or oppose a challenge to the confidentiality designation) at a later date. Any challenge

to the confidentiality designation of the Training Data or portions thereof under this Training Data

Inspection Protocol shall be written, shall be served on outside counsel for OpenAI, shall

particularly identify the documents or information that the Inspecting Party contends should be

differently designated, and shall state the grounds for the objection. The parties shall meet and

confer in a good faith effort to resolve the dispute. Notwithstanding any challenge to a designation,

the Training Data in question shall continue to be treated as "HIGHLY CONFIDENTIAL –

ATTORNEYS' EYES ONLY" until one of the following occurs: (1) OpenAI withdraws such

designation in writing; or (2) the Court rules that the Training Data in question is not entitled to

the designation.

6.   Nothing in this Training Data Inspection Protocol shall alter or change in any way

the requirements of the Stipulated Protective Order. In the event of any conflict, however, this

Training Data Inspection Protocol shall control for any Training Data made available for

inspection.

7.   Training Data shall be produced for inspection and review subject to the following

provisions:

a.   Training Data shall be made available by OpenAI in a secure room on a

secured computer without Internet access or network access to other unauthorized computers or

devices. The secured computer will contain a README file that will provide a directory of the

Training Data and brief descriptions of layout, format, and searching.

b.   The secured computer will be equipped with tools that are sufficient for

viewing and searching the Training Data made available for inspection. Defendants will

reasonably cooperate with Plaintiffs to address any technical concerns or objections Plaintiffs may

have regarding the hardware and software that is provided to conduct the Training Data review. Plaintiffs reserve all rights to seek any additional relief from the Court, including to enable a more efficient and/or effective review of the Training Data.

        c.   The Producing Party shall provide the receiving Party with information explaining how to start, log on to, and operate the secured computer(s) in order to access the Training Data on the secured computer(s). The Producing Party's outside counsel will be available electronically to make reasonable efforts to attempt to resolve issues that may arise during the course of inspection.

        d.   The Inspecting Party's counsel and/or experts may request that software tools and/or files be installed on the secured computer, provided, however, that (a) the Inspecting Party possesses an appropriate license to such software tools and/or files; (b) OpenAI approves such software tools and/or files, such approval not to be unreasonably withheld; and (c) such other software tools and/or files are reasonably necessary for the Inspecting Party to perform its review of the Training Data consistent with all of the protections herein. The Inspecting Party must provide OpenAI with the licensed software tool(s) and/or files, at the Inspecting Party's expense, at least seven days in advance of the date upon which the Inspecting Party wishes to have the additional software tools and/or files available for use on the secured computer. The Producing Party will install and confirm installation of said software on the Source Code Computers prior to the inspection.

        e.   No recordable media or recordable devices, including without limitation computers, cellular telephones, cameras, other recording devices, or drives of any kind, shall be permitted into the secure inspection room, except the Producing Party may provide a limited-use note-taking computer at Inspecting Party's request, solely for note-taking purposes. At the end of

each day of inspection, the Inspecting Party shall be able to copy notes from the note taking

computer onto a recordable device, under the supervision of the Producing Party.

        f.   The Inspecting Party's counsel and/or experts may take handwritten notes

or electronic notes on the note-taking computer in scratch files, but may not copy any Training

Data itself into any notes. For the avoidance of doubt, this provision shall not prevent the

Inspecting Party's counsel and/or experts from recording in their notes particular items, files, or

categories of items or files contained in the Training Data. The Inspecting Party will not waive any

applicable work-product protection over their electronic notes by saving them to the note-taking

computer temporarily. Such notes may not be in encoded or encrypted form. Any notes related to

the Training Data will be treated as "HIGHLY CONFIDENTIAL – ATTORNEYS' EYES

ONLY."

        g.   The Producing Party may visually monitor the activities of the Inspecting

Party's representatives during any inspection, but only to ensure that there is no unauthorized

recording, copying, or transmission of the Training Data. Any monitoring must be conducted from

outside the room where the inspection is taking place.

        h.   No copies of all or any portion of the Training Data, or other written or

electronic record of the Training Data, may leave the secured room in which the Training Data is

inspected except as provided herein. The Inspecting Party may obtain print outs of limited portions

of the Training Data or electronic notes taken on the note-taking computer to prepare court filings

or pleadings or other papers (including a testifying expert's expert report) by following the

procedures provided herein. For purposes of this protocol, references to "print," "printing," or

"print outs" are understood to refer to a Bates-stamped electronic production (as described in this

Paragraph). To make a request, the Inspecting Party shall create a directory entitled "Print

Request" and save the desired limited portions of the Training Data or notes in that directory. The

beginning of each portion of Training Data the Inspecting Party wishes to print must include the

filename, file path, and line numbers where the material was found in the training data or other

information that allows for specific identification of the material. The Inspecting Party shall alert

OpenAI when it has saved the desired limited portions of the Training Data or notes in the "Print

Request" directory that it requests to be printed. Upon receiving a request, OpenAI shall Bates

number, and label 'HIGHLY CONFIDENTIAL – ATTORNEYS' EYES ONLY" all requested

pages. Within 7 business days from the date of request, OpenAI shall either (i) produce electronic

versions to the Inspecting Party's counsel, or (ii) inform the Inspecting Party that OpenAI objects

that the requested portions are excessive, not for a permitted purpose, and/or not justified (*see,

e.g.* Fed. R. Civ. Pro. 26(b)). In the event that OpenAI objects, the parties shall meet and confer

within five business days of OpenAI's notice of its objection. If, after meeting and conferring,

OpenAI and the Inspecting Party cannot resolve the objection, the Inspecting Party shall be entitled

to seek a Court resolution of whether the requested Training Data should be produced.

        i.   All persons who will review OpenAI's Training Data on behalf of an

Inspecting Party, including the Inspecting Party's counsel, must qualify under paragraph 16 of the

Stipulated Protective Order as an individual to whom "HIGHLY CONFIDENTIAL –

ATTORNEYS' EYES ONLY" information may be disclosed, and must sign the Non-Disclosure

Agreement attached as Exhibit A to the Stipulated Protective Order. All persons who review

OpenAI's Training Data in the secured inspection room or on the secured computer on behalf of

an Inspecting Party shall also be identified in writing to OpenAI at least seven days in advance of

the first time that such person reviews such Training Data. All authorized persons viewing

Training Data in the secured inspection room or on the secured computer shall, on each day they

view Training Data, sign a log that will include the names of persons who enter the locked room to view the Training Data and when they enter and depart. Proper identification of all authorized persons shall be provided prior to any access to the secure inspection room or the secured computer containing Training Data. Proper identification requires showing, at a minimum, a photo identification card sanctioned by the government of any State of the United States, by the government of the United States, or by the nation state of the authorized person's current citizenship. Access to the secure inspection room or the secured computer may be denied, at the discretion of OpenAI, to any individual who fails to provide proper identification.

j.   Unless otherwise agreed in advance by the parties in writing, following each day on which inspection is done under this protocol, the Inspecting Party's counsel and/or experts shall remove all notes, documents, and all other materials from the secure inspection room. OpenAI shall not be responsible for any items left in the room following each inspection session, and the Inspecting Party shall have no expectation of confidentiality for any items left in the room following each inspection session without a prior agreement to that effect.

k.   Other than as provided above, the Inspecting Party will not copy, remove, or otherwise transfer any Training Data from the secured computer including, without limitation, copying, removing, or transferring the Training Data onto any recordable media or recordable device. The Inspecting Party will not transmit any Training Data in any way from OpenAI's facilities.

l.   Notwithstanding any provisions of this Training Data Protocol or the Protective Order entered on May 16, 2024, the Parties reserve the right to amend this protocol either by written agreement or Order of the Court upon showing of good cause.

IT IS SO STIPULATED, through Counsel of Record.

Dated: September 5, 2024 */s/ Alejandra C. Salinas* Alejandra C. Salinas (*pro hac vice*) SUSMAN GODFREY L.L.P. 1000 Louisiana Street, Suite 5100 Houston, TX 77002 Telephone: 713-651-9366 asalinas@susmangodfrey.com ***Interim Co-Lead Class Counsel***	Dated: September 5, 2024 /s/ Thomas E. Gorman Thomas E. Gorman KEKER, VAN NEST & PETERS LLP 633 Battery Street San Francisco, CA 94111-1809 Telephone: 415 391 5400 tgorman@keker.com ***Attorneys for Defendants OPENAI INC., OPENAI LP, OPENAI LLC, OPENAI GP LLC, OPENAI OPCO LLC, OPENAI GLOBAL LLC, OAI CORPORATION LLC, OPENAI HOLDINGS LLC, OPENAI STARTUP FUND I LP, OPENAI STARTUP FUND GP I LLC, and OPENAI STARTUP FUND MANAGEMENT LLC***

Date: _____   SO ORDERED:

_____
ONA T. WANG
United States Magistrate Judge